## **PCT**

# WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6:

C12N 15/82, 15/29, 15/53, A01H 5/00,
5/10

(11) International Publication Number: WO 98/45460

(43) International Publication Date: 15 October 1998 (15.10.98)

(21) International Application Number: PCT/US98/07178
 (22) International Filing Date: 9 April 1998 (09.04.98)

08/831,570 9 April 1997 (09.04.97) US

(71) Applicant (for all designated States except US):

(71) Applicant (for all designated States except US): RHONE-POULENC AGRO [FR/FR]; Dépt. Propriété Industrielle, 14-20, rue Pierre Baizet, F-69009 Lyon (FR).

(72) Inventors; and
(75) Inventors/Applicants (for US only): THOMAS, Terry, L. [US/US]; 2804 Cloister Drive, College Station, TX 77845 (US). BEREMAND, Phillip, D. [US/US]; 9208 Brookwater Circle, College Station, TX 77845 (US). NUNBERG,

ter Circle, College Station, TX 77845 (US). NUNBERG, Andrew, N. [US/US]; 12215-B Encanto Lane, Maryland Heights, MO 63043 (US).

(74) Agents: DiGIGLIO, Frank, S. et al.; Scully, Scott, Murphy & Presser, 400 Garden City Plaza, Garden City, NY 11530 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

#### **Published**

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: A SUNFLOWER ALBUMIN 5' REGULATORY REGION FOR THE MODIFICATION OF PLANT SEED LIPID COMPOSITION

#### (57) Abstract

(30) Priority Data:

The present invention is directed to 5' regulatory regions of a sunflower albumin gene. The 5' regulatory regions, when operably linked to either the coding sequence of a heterologous gene or a sequence complementary to a native plant gene direct expression of the coding sequence or complementary sequence in a plant seed. The regulatory regions are useful in expression cassettes and expression vectors for the transformation of plants. Also provided are methods of modulating the levels of a heterologous gene or native plant gene such as a fatty acid synthesis or lipid metabolism gene by transforming a plant with the subject expression cassettes and expression vectors.

## FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon .	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of Americ
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
Cl	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

# A SUNFLOWER ALBUMIN 5' REGULATORY REGION FOR THE MODIFICATION OF PLANT SEED LIPID COMPOSITION

### BACKGROUND OF THE INVENTION

Seed oil content has traditionally been modified by plant breeding. The use of recombinant DNA technology to alter seed oil composition can accelerate this process and in some cases alter seed oils in a way that cannot be accomplished by breeding alone. The oil composition of Brassica has been significantly altered by modifying the expression of a number of lipid metabolism genes. Such manipulations of seed oil composition have focused on altering the proportion of endogenous component fatty acids. For example, antisense repression of the Δ12-desaturase gene in transgenic rapeseed has resulted in an increase in oleic acid of up to 83%. Topfer et al. 1995 Science 268:681-686.

There have been some successful attempts at modifying the composition of seed oil in transgenic plants by introducing new genes that allow the production of a fatty acid that the host plants were not previously capable of synthesizing. Van de Loo, et al. (1995 Proc. Natl. Acad. Sci USA 92:6743-6747) have been able to introduce a Δ12-hydroxylase gene into transgenic tobacco, resulting in the introduction of a novel fatty acid, ricinoleic acid, into its seed oil. The reported accumulation was modest from plants carrying constructs in which transcription of the hydroxylase gene was under the control of the

-2-

cauliflower mosaic virus (CaMV) 35S promoter.

Similarly, tobacco plants have been engineered to produce low levels of petroselinic acid by expression of an acyl-ACP desaturase from coriander (Cahoon et al. 1992 Proc. Natl. Acad. Sci USA 89:11184-11188).

The long chain fatty acids (C18 and larger), have significant economic value both as nutritionally and medically important foods and as industrial commodities (Ohlrogge, J.B. 1994 Plant Physiol. 104:821-826). Linoleic (18:2  $\Delta 9,12$ ) and  $\alpha$ -linolenic

acid (18:3 Δ9,12,15) are essential fatty acids found in many seed oils. The levels of these fatty-acids have been manipulated in oil seed crops through breeding and biotechnology (Ohlrogge, et al. 1991 Biochim. Biophys. Acta 1082:1-26; Topfer et al. 1995

Science 268:681-686). Additionally, the production of novel fatty acids in seed oils can be of considerable use in both human health and industrial applications.

Consumption of plant oils rich in γ-linolenic acid (GLA) (18:3 Δ6,9,12) is thought to alleviate hypercholesterolemia and other related clinical disorders which correlate with susceptibility to coronary heart disease (Brenner R.R. 1976 Adv. Exp. Med. Biol. 83:85-101). The therapeutic benefits of dietary GLA may result from its role as a precursor to prostaglandin synthesis (Weete, J.D. 1980 in Lipid Biochemistry of Fungi and Other Organisms, eds. Plenum Press, New York, pp. 59-62). Linoleic acid(18:2) (LA) is transformed into gamma linolenic acid (18:3) (GLA) by the enzyme Δ6-desaturase.

30

- 3 -

Few seed oils contain GLA despite high 1 contents of the precursor linoleic acid. This is due to the absence of  $\Delta 6$ -desaturase activity in most plants. For example, only borage (Borago officinalis), evening primrose (Oenothera biennis), and currants (Ribes nigrum) produce appreciable amounts of linolenic acid. Of these three species, only Oenothera and borage are cultivated as a commercial source for GLA. It would be beneficial if agronomic seed oils could be engineered to produce GLA in significant quantities by introducing a heterologous A6-desaturase gene. It would also be beneficial if other expression products associated with fatty acid synthesis and lipid metabolism could be produced in plants at high enough levels so that commercial production of a particular expression product becomes feasible.

As disclosed in U.S. Patent No. 5,552,306, a cyanobacterial Δ<sup>6</sup>-desaturase gene has been recently isolated. Expression of this cyanobacterial gene in transgenic tobacco resulted in significant but low level GLA accumulation. (Reddy et al. 1996 Nature Biotech. 14:639-642). Applicant's copending U.S. Application Serial No. 08,366,779, discloses a Δ6-desaturase gene isolated from the plant Borago officinalis and its expression in tobacco under the control of the CaMV 35S promoter. Such expression resulted in significant but low level GLA and octadecatetraenoic acid (ODTA or OTA) accumulation in seeds. Thus, a need exists for a promoter which functions in plants and which consistently directs

-4-

high level expression of lipid metabolism genes in transgenic plant seeds.

Sunflower embryos accumulate two major classes of storage proteins. These are the 11 S globulins, soluble in 1 M NaC1, and 2 S albumins, soluble in water (Youle et al. 1981 Am J. Bot 68:44-The synthesis, processing and accumulation of 2 S albumin seed proteins have been studied intensively in Brassica napus (Crouch et al., 1983 J. Mol. Appl. Genet. 2:273-284; Ericson et al., 1986 J. Biol. Chem. 261:14576-14581), pea (Higgins et al., 1986 Plant Mol. Biol. 8:37-45), radish (Laroche-Raynal et al., 1986 Eur. J. Biochem. 157:321-327), castor bean (Lord J.M., 1985 Eur. J: Biochem 146:403-409) and Brazil nut (Sun et al., 1987 Eur. J. Biochem 162:477-483). A major conclusion of these studies is that the characteristic low molecular weight, disulfide-linked albumin polypeptides found in mature seeds result from the extensive processing of larger precursors synthesized during embryogenesis. Two additional characteristics 20 that define the 2 S albumin seed storage proteins are high amide content and high frequency of cysteine residues (Youle et al., 1981).

In sunflower, the 2 S albumins represent
more than 50% of the protein present in seeds (Youle
et al., 1981) and consist of two or three closely
related polypeptides with molecular weights of
approximately 19 kDa (Cohen, E.A., 1986 "Analysis of
sunflower 2S seed storage protein genes" MS thesis,
Texas A&M University; Allen et al. 1987 Plant Mol Biol
5:165-173). The sunflower albumin is apparently

- 5 -

maintained in a compact structure by intramolecular disulfide bonds resulting in a rapidly migrating species with an apparent molecular weight of 14 kDa when analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) under non-reducing

conditions. When reduced, this species migrates as a 19 kDa polypeptide (Cohen, E.A., 1986). In contrast, most other 2 S proteins are composed of large and small subunit polypeptides, derived from a single precursor, and linked by itnermoleclar disulfide bonds

10 (Crouch et al. 1983 J. Mol. Appl. Genet. 2:273-284; Ericson et al. 1986 J. Biol. Chem. 261:14576-14581; Sun et al. 1987, Eur. J. Bioch. 162:477-483.)

Albumin polypeptides can be detected in sunflower embryos by 5 days post-fertilization (DPF), 2 days before helianthinins are detectable, and continue to accumulate through seed maturation. Sunflower albumin mRNAs, also first detected at 5 DPF, accumulate rapidly in sunflower embryos reaching maximum prevalence between 12 and 15 DPF. After this time albumin transcripts decrease in prevalence with kinetics similar to that observed for helianthinin mRNA (Allen et al. 1987). Functional sunflower albumin mRNAs are undetectable in dry seeds, germinated seedlings or leaves (Cohen 1986).

A number of albumin cDNAs and genomic clones have been isolated from different plant species including sunflower (Allen et al. 1987 Mol-Gen Genet. 210:211-218) and pea (Higgins et al. 1986 J. Biol. Chem 261:11124-11130). As in other classes of seed proteins such as Brassica napis (Crouch et al., 1983;

-6-

Ericson et al., 1986), 2 S albumin seed proteins are encoded by small gene families.

The present invention provides 5' regulatory sequences from a sunflower albumin gene which direct high level expression of lipid metabolism genes in transgenic plants. In accordance with the present invention, chimeric constructs comprising a sunflower albumin 5' regulatory region operably linked to coding sequence for a lipid metabolism gene such as a Δ6-desaturase gene are provided. Transgenic plants comprising the subject chimeric constructs accumulate GLA to approximately 10% of C18 fatty acids. This is within the range of accumulation of GLA for Oenothera biennis, a primary commercial source for GLA.

## 15 SUMMARY OF THE INVENTION

The present invention is directed to 5' regulatory regions of a sunflower albumin gene. The 5' regulatory regions, when operably linked to either the coding sequence of a heterologous gene or sequence complementary to a native plant gene, direct expression of the heterologous gene or complementary sequence in a plant seed.

The present invention thus provides expression cassettes and expression vectors comprising an albumin 5' regulatory region operably linked to a heterologous gene or a sequence complementary to a native plant gene.

Plant transformation vectors comprising the expression cassettes and expression vectors are also provided as are plant cells transformed by these

-7-

vectors, and plants and their progeny containing the  $^{\mbox{\scriptsize l}}$  vectors.

In one embodiment of the invention, the heterologous gene or complementary sequence is a fatty acid synthesis gene or a lipid metabolism gene.

In another aspect of the present invention, a method is provided for producing a plant with increased levels of a product of a fatty acid synthesis or lipid metabolism gene.

In particular, there is provided a method

for producing a plant with increased levels of a fatty
acid synthesis or lipid metabolism gene by
transforming a plant with the subject expression
cassettes and expression vectors which comprise an
albumin 5' regulatory region and a coding sequence for
a fatty acid synthesis or lipid metabolism gene.

In another aspect of the present invention, there is provided a method for cosuppressing a native fatty acid synthesis or lipid metabolism gene by transforming a plant with the subject expression cassettes and expression vectors which comprise an albumin 5' regulatory region and a coding sequence for a fatty acid synthesis or lipid metabolism gene.

A further aspect of this invention provides a method of decreasing production of a native plant gene such as a fatty acid synthesis gene or a lipid metabolism gene by transforming a plant with an expression vector comprising an albumin 5' regulatory region operably linked to a nucleic acid sequence complementary to a native plant gene.

-8-

Also provided are methods of modulating the levels of a heterologous gene or native plant gene such as a fatty acid synthesis or lipid metabolism gene by transforming a plant with the subject expression cassettes and expression vectors.

5

## BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 depicts the nucleotide and corresponding amino acid sequence of the borage Δ6-desaturase gene (SEQ ID NO:1). The cytochrome b5 10 heme-binding motif is boxed and the putative metal binding, histidine rich motifs (HRMs) are underlined. The motifs recognized by the primers (PCR analysis) are underlined with dotted lines, i.e. tgg aaa tgg aac cat aa; and gag cat cat ttg ttt cc.

15 Fig. 2 is a dendrogram showing similarity of the borage Δ6-desaturase to other membrane-bound desaturases. The amino acid sequence of the borage Δ6-desaturase was compared to other known desaturases using Gene Works (IntelliGenetics). Numerical values correlate to relative phylogenetic distances between subgroups compared.

Fig. 3A provides a gas liquid chromatography profile of the fatty acid methyl esters (FAMES) derived from leaf tissue of a wild type tobacco 'Xanthi'.

Fig. 3B provides a gas liquid chromatography profile of the FAMES derived from leaf tissue of a tobacco plant transformed with the borage  $\Delta 6$ -desaturase cDNA under transcriptional control of the 30 CaMV 35S promoter (pAN2). Peaks corresponding to

-9-

methyl linoleate (18:2), methyl  $\gamma$ -linolenate (18:3 $\gamma$ ), methyl  $\alpha$ -linolenate (18:3 $\alpha$ ), and methyl octadecatetraenoate (18:4) are indicated.

Fig. 4 is the nucleotide sequence of the HaG5 regulatory region. The transcriptional start site (+1) is indicated by a bold T. The underlined Bam HI restriction site was introduced by PCR.

Fig. 5 is a scheme depicting construction of the sunflower albumin HaG5 regulatory region/ $\Delta 6$ -desaturase gene expression vector.

Fig. 6A is an RNA gel blot analysis carried out on 5  $\mu$ g samples of RNA isolated from borage leaf, root, and 12 dpp embryo tissue, using labeled borage  $\Delta 6$ -desaturase cDNA as a hybridization probe.

Fig. 6B depicts a graph corresponding to the Northern analysis results for the experiment shown in Fig. 6A.

Fig. 7 is a PCR analysis showing the presence of the borage delta 6-desaturase gene in transformed plants of oilseed rape. Lanes 1,3 and 4 were loaded with PCR reactions performed with DNA from plants transformed with the borage delta 6-desaturase gene linked to the oleosin 5' regulatory region; lane 2: DNA from plant transformed with the borage delta 6-desaturase gene linked to the albumin 5' regulatory region; lanes 5 and 6: DNA from non-transformed plants; lane 7: molecular weight marker (1 kb ladder, Gibco BRL); lane 8: PCR without added template DNA; lane 9: control with DNA from Agrobacterium tumefaciens EHA 105 containing the plasmid pAN3 the

-10-

borage delta 6-desaturase gene linked to the oleosin 1 5' regulatory region.

### DETAILED DESCRIPTION OF THE INVENTION

The present invention provides isolated 5 nucleic acids encoding 5' regulatory regions from a sunflower albumin gene. In accordance with the present invention, the subject 5' regulatory regions, when operably linked to either a coding sequence of a heterologous gene or sequence complementary to a native plant gene, direct expression of the coding sequence or the complementary sequence in a plant seed. The albumin 5' regulatory regions of the present invention are useful in the construction of an expression cassette which comprises in the 5' to 3' 15 direction, a subject albumin 5' regulatory region, a heterologous gene or sequence complementary to a native plant gene under control of the regulatory region and a 3' termination sequence. Such an expression cassette can be incorporated into a variety 20 of autonomously replicating vectors in order to construct an expression vector.

In accordance with the present invention, it has been surprisingly found that plants transformed with a subject expression vector accumulate GLA to approximately 10% of C18 fatty acids. Such an accumulation is within the range of accumulation of GLA for Oenothera biennis, a primary commercial source for GLA.

As used herein, the term "cassette" refers 30 to a nucleotide sequence capable of expressing a

-11-

particular gene if said gene is inserted so as to be 1 operably linked to one or more regulatory regions present in the nucleotide sequence. Thus, for example, the expression cassette may comprise a heterologous coding sequence which is desired to be 5 expressed in a plant seed. The expression cassettes and expression vectors of the present invention are therefore useful for directing seed-specific expression of any number of heterologous genes. term "seed-specific expression" as used herein, refers  $^{10}$  to expression in various portions of a plant seed such as the endosperm and embryo.

An isolated nucleic acid encoding a 5' regulatory region from a sunflower albumin gene can be provided as follows. Albumin recombinant genomic  $^{15}$  clones are isolated by screening a sunflower genomic DNA library with a cDNA (or a portion thereof) representing albumin mRNA. A number of different albumin cDNAs have been isolated. The methods used to isolate such cDNAs as well as the nucleotide and 20 corresponding amino acid sequences have been published. Higgins et al., 1986 J. Biol. Chem. 261: 11124-11130; Allen et al., 1987 in Molecular Approaches to Developmental Biology, Alan R. Liss, Inc., pp. 415-424.

25 Methods considered useful in obtaining albumin genomic recombinant DNA are provided in Sambrook et al. 1989, in Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY, for example, or any of the myriad of laboratory manuals on

 $^{
m 30}$  recombinant DNA technology that are widely available.

-12-

To determine nucleotide sequences, a multitude of techniques are available and known to the ordinarily skilled artisan. For example, restriction fragments containing an albumin regulatory region can be subcloned into the polylinker site of a sequencing vector such as pBluescript (Stratagene). These pBluescript subclones can then be sequenced by the double-stranded dideoxy method (Chen and Seeburg, 1985, DNA 4:165).

In a preferred embodiment, the sunflower

albumin regulatory region comprises nucleotides 860 to
+29 of Fig. 4 (nucleotides 1-895 of SEQ ID NO:2).

Modifications to the albumin regulatory region as set
forth in SEQ ID NO:2 which maintain the characteristic
property of directing seed-specific expression, are

within the scope of the present invention. Such
modifications include insertions, deletions and
substitutions of one or more nucleotides.

The 5' regulatory region of the present invention can be derived from restriction endonuclease or exonuclease digestion of an albumin genomic clone. Thus, for example, the known nucleotide or amino acid sequence of the coding region of an isolated albumin gene is aligned to the nucleic acid or deduced amino acid sequence of an isolated albumin genomic clone and 5' flanking sequence (i.e., sequence upstream from the translational start codon of the coding region) of the isolated albumin genomic clone located.

The albumin 5' regulatory region as set forth in SEQ ID NO:2 (nucleotides -860 to +29 of Fig. 30 4) may be generated from a genomic clone having either

-13-

or both excess 5' flanking sequence or coding sequence

by exonuclease III-mediated deletion. This is
accomplished by digesting appropriately prepared DNA
with exonuclease III (exoIII) and removing aliquots at
increasing intervals of time during the digestion.

The resulting successively smaller fragments of DNA may be sequenced to determine the exact endpoint of the deletions. There are several commercially available systems which use exonuclease III (exoIII) to create such a deletion series, e.g. Promega

Biotech, "Erase-A-Base" system. Alternatively, PCR primers can be defined to allow direct amplification of the subject 5' regulatory regions.

Using the same methodologies, the ordinarily skilled artisan can generate one or more deletion fragments of nucleotides 1-895 as set forth in SEQ ID NO:2. Any and all deletion fragments which comprise a contiguous portion of nucleotides set forth in SEQ ID NO:2 and which retain the capacity to direct seed-specific expression are contemplated by the present invention.

The identification of albumin 5' regulatory sequences which direct seed-specific expression comprising nucleotides 1-895 of SEQ ID NO:2 and modifications or deletion fragments thereof, can be accomplished by transcriptional fusions of specific sequences with the coding sequences of a heterologous gene, transfer of the chimeric gene into an appropriate host, and detection of the expression of the heterologous gene. The assay used to detect expression depends upon the nature of the heterologous

-14-

sequence. For example, reporter genes, exemplified by chloramphenical acetyl transferase and  $\beta$ -glucuronidase (GUS), are commonly used to assess transcriptional and translational competence of chimeric constructions. Standard assays are available to sensitively detect

- the reporter enzyme in a transgenic organism. The β-glucuronidase (GUS) gene is useful as a reporter of promoter activity in transgenic plants because of the high stability of the enzyme in plant cells, the lack of intrinsic β-glucuronidase activity in higher plants
- and availability of a quantitative fluorimetric assay and a histochemical localization technique.

  Jerfferson et al. (1987 EMBO J 6:3901) have established standard procedures for biochemical and histochemical detection of GUS activity in plant
- tissues. Biochemical assays are performed by mixing plant tissue lysates with 4-methylumbelliferyl- $\beta$ -D-glucuronide, a fluorimetric substrate for GUS, incubating one hour at 37°C, and then measuring the fluorescence of the resulting 4-methyl-umbelliferone.
- Histochemical localization for GUS activity is determined by incubating plant tissue samples in 5-bromo-4-chloro-3-indolyl-glucuronide (X-Gluc) for about 18 hours at 37°C and observing the staining pattern of X-Gluc. The construction of such chimeric
- genes allows definition of specific regulatory sequences and demonstrates that these sequences can direct expression of heterologous genes in a seed-specific manner.

Another aspect of the invention is directed 30 to a chimeric plant gene comprising a 5' regulatory

-15-

region from an albumin gene which directs seed

specific expression operably linked to the coding sequence of a heterologous gene such that the regulatory element is capable of controlling expression of the product encoded by the heterologous gene. The heterologous gene can be any gene other than albumin. If necessary, additional regulatory elements or parts of these elements sufficient to cause expression resulting in production of an effective amount of the polypeptide encoded by the heterologous gene are included in the chimeric constructs.

Accordingly, the present invention provides chimeric genes comprising sequences of the albumin 5' regulatory region that confer seed-specific expression 15 which are operably linked to a sequence encoding a heterologous gene such as a lipid metabolism enzyme. Examples of lipid metabolism and fatty acid synthesis genes useful for practicing the present invention include lipid desaturases such as A6-desaturases, A12desaturases,  $\Delta 15$ -desaturases and other related desaturases such as stearoyl-ACP desaturases, acyl carrier proteins (ACPs), thioesterases, acetyl transacylases, acetyl-coA carboxylases, ketoacylsynthases, malonyl transacylases, and elongases. Such lipid metabolism and fatty acid synthesis genes have been isolated and characterized from a number of different bacteria and plant species. Their nucleotide coding sequences as well as methods of isolating such coding sequences are disclosed in the

-16-

published literature and are widely available to those  $\ensuremath{^{1}}$  of skill in the art.

In particular, the \( \Delta 6 \)-desaturase genes disclosed in U.S. Patent No. 5,552,306 and applicants' copending U.S. Application Serial No. 08/366,779 filed December 30, 1994 and incorporated herein by reference, are contemplated as lipid metabolism genes particularly useful in the practice of the present invention.

The chimeric genes of the present invention are constructed by ligating a 5' regulatory region of an albumin genomic DNA to the coding sequence of a heterologous gene. The juxtaposition of these sequences can be accomplished in a variety of ways. In a preferred embodiment the order of the sequences, from 5' to 3', is an albumin 5' regulatory region (including a promoter), a coding sequence, and a termination sequence which includes a polyadenylation site.

chimeric genes are well known to those of ordinary skill in the art and can be found in references such as Sambrook et al.(1989). A variety of strategies are available for ligating fragments of DNA, the choice of which depends on the nature of the termini of the DNA fragments. One of ordinary skill in the art recognizes that in order for the heterologous gene to be expressed, the construction requires promoter elements and signals for efficient polyadenylation of the transcript. Accordingly, the albumin 5' regulatory region that contains the consensus promoter

-17-

sequence known as the TATA box can be ligated directly to a promoterless heterologous coding sequence.

The restriction or deletion fragments that contain the albumin TATA box are ligated in a forward orientation to a promoterless heterologous gene such as the coding sequence of  $\beta$ -glucuronidase (GUS). The skilled artisan will recognize that the subject albumin 5' regulatory regions can be provided by other means, for example chemical or enzymatic synthesis. The 3' end of a heterologous coding sequence is

optionally ligated to a termination sequence comprising a polyadenylation site, exemplified by, but not limited to, the nopaline synthase polyadenylation site, or the octopine T-DNA gene 7 polyadenylation site. Alternatively, the polyadenylation site can be provided by the heterologous gene.

The present invention also provides methods of increasing levels of heterologous genes in plant seeds. In accordance with such methods, the subject expression cassettes and expression vectors are introduced into a plant in order to effect expression of a heterologous gene. For example, a method of producing a plant with increased levels of a product of a fatty acid synthase or lipid metabolism gene is provided by transforming a plant cell with an expression vector comprising an albumin 5' regulatory region operably linked to a fatty acid synthesis or lipid metabolism gene and regenerating a plant with increased levels of the product of said fatty acid synthesis or lipid metabolism gene.

-18-

Another aspect of the present invention l provides methods of reducing levels of a product of a gene which is native to a plant which comprises transforming a plant cell with an expression vector comprising a subject albumin regulatory region operably linked to a nucleic acid sequence which is complementary to the native plant gene. In this manner, levels of endogenous product of the native plant gene are reduced through the mechanism known as antisense regulation. Thus, for example, levels of a product of a fatty acid synthesis gene or lipid metabolism gene are reduced by transforming a plant with an expression vector comprising a subject albumin 5' regulatory region operably linked to a nucleic acid sequence which is complementary to a nucleic acid sequence coding for a fatty acid synthesis or lipid metabolism gene.

The present invention also provides a method of cosuppressing a gene which is native to a plant which comprises transforming a plant cell with an expression vector comprising a subject albumin regulatory region operably linked to a nucleic acid sequence coding for the native plant gene. manner, levels of endogenous product of the native plant gene are reduced through the mechanism known as cosuppression. Thus, for example, levels of a product of a fatty acid synthesis gene or lipid metabolism gene are reduced by transforming a plant with an expression vector comprising a subject albumin 5' regulatory region operably linked to a nucleic acid · 30 sequence coding for a fatty acid synthesis or lipid

20

- 19 -

metabolism gene native to the plant. Although the exact mechanism of cosuppression is not completely understood, one skilled in the art is familiar with published works reporting the experimental conditions and results associated with cosuppression (Napoli et al. 1990 The Plant Cell 2:270-289; Van der Krol 1990 The Plant Cell 2:291-299.

To provide regulated expression of the heterologous or native genes, plants are transformed with the chimeric gene constructions of the invention. 10 Methods of gene transfer are well known in the art. The chimeric genes can be introduced into plants by leaf disk transformation-regeneration procedure as described by Horsch et al. (1985) Science 227:1229. Other methods of transformation such as protoplast culture (Horsch et al. 1984 Science 223:496, DeBlock et al. 1984 EMBO J. 2:2143, Barton et al. 1983, Cell 32:1033) can also be used and are within the scope of this invention. In a preferred embodiment, plants are transformed with Agrobacterium-derived vectors such as 20 those described in Klett et al. (1987) Annu. Rev. Plant Physiol. 38:467. Other well-known methods are available to insert the chimeric genes of the present invention into plant cells. Such alternative methods include biolistic approaches (Klein et al. 1987 Nature 327:70), electroporation, chemically-induced DNA uptake, and use of viruses or pollen as vectors. When necessary for the transformation

When necessary for the transformation method, the chimeric genes of the present invention can be inserted into a plant transformation vector, e.g. the binary vector described by Bevan, M. 1984

-20-

Nucl. Acids Res. 12:8711-8721. Plant transformation l vectors can be derived by modifying the natural gene transfer system of Agrobacterium tumefaciens. natural system comprises large Ti (tumor-inducing) plasmids containing a large segment, known as T-DNA, which is transferred to transformed plants. Another segment of the Ti plasmid, the vir region, is responsible for T-DNA transfer. The T-DNA region is bordered by terminal repeats. In the modified binary vectors, the tumor inducing genes have been deleted 10 and the functions of the vir region are utilized to transfer foreign DNA bordered by the T-DNA border sequences. The T-region also contains a selectable marker for antibiotic resistance, and a multiple cloning site for inserting sequences for transfer. 15 Such engineered strains are known as "disarmed" A. tumefaciens strains, and allow the efficient transfer of sequences bordered by the T-region into the nuclear

Surface-sterilized leaf disks or other

susceptible tissues are inoculated with the "disarmed" foreign DNA-containing A. tumefaciens, cultured for a number of days, and then transferred to antibiotic-containing medium. Transformed shoots are then selected after rooting in medium containing the appropriate antibiotic, and transferred to soil. Transgenic plants are pollinated and seeds from these plants are collected and grown on antibiotic medium.

genome of plants.

Expression of a heterologous or reporter gene in developing seeds, young seedlings and mature plants can be monitored by immunological,

-21-

histochemical or activity assays. As discussed herein,

the choice of an assay for expression of the chimeric
gene depends upon the nature of the heterologous
coding region. For example, Northern analysis can be
used to assess transcription if appropriate nucleotide
probes are available. If antibodies to the
polypeptide encoded by the heterologous gene are
available, Western analysis and immunohistochemical
localization can be used to assess the production and
localization of the polypeptide. Depending upon the
heterologous gene, appropriate biochemical assays can
be used. For example, acetyltransferases are detected
by measuring acetylation of a standard substrate. The
expression of a lipid desaturase gene can be assayed
by analysis of fatty acid methyl esters (FAMES).

15 Another aspect of the present invention provides transgenic plants or progeny of these plants containing the chimeric genes of the invention. Both monocotyledonous and dicotyledonous plants are contemplated. Plant cells are transformed with the 20 chimeric genes by any of the plant transformation methods described above. The transformed plant cell, usually in the form of a callus culture, leaf disk or whole plant (via the vacuum infiltration method of Bechtold et al. 1993 C.R. Acad. Sci. Paris, 316:1194- $^{25}$  1199) is regenerated into a complete transgenic plant by methods well-known to one of ordinary skill in the art (e.g. Horsh et al. 1985 Science 227:1129). preferred embodiment, the transgenic plant is sunflower, cotton, oil seed rape, maize, tobacco, 30 Arabidopsis, peanut or soybean. Since progeny of

-22-

transformed plants inherit the chimeric genes, seeds

or cuttings from transformed plants are used to
maintain the transgenic line.

The following examples further illustrate the invention.

-23-

#### EXAMPLE 1

Isolation of Membrane-Bound Polysomal
RNA and Construction of Borage cDNA Library

Membrane-bound polysomes were isolated from borage seeds 12 days post pollination (12 DPP) using the protocol established for peas by Larkins and Davies (1975 Plant Phys. 55: 749-756). RNA was extracted from the polysomes as described by Mechler (1987 Methods in Enzymology 152: 241-248, Academic Press). Poly-A+ RNA was isolated from the membrane bound polysomal RNA using Oligotex-dTM beads (Qiagen).

Corresponding cDNA was made using
Stratagene's ZAP cDNA synthesis kit. The cDNA library
was constructed in the lambda ZAP II vector
(Stratagene) using the lambda ZAP II kit. The primary
library was packaged with Gigapack II Gold packaging
extract (Stratagene).

20

25

-24-

#### EXAMPLE 2

1 Isolation of a  $\Delta$ -6 Desaturase cDNA from Borage

#### Hybridization protocol

5

The amplified borage cDNA library was plated at low density (500 pfu on 150 mm petri dishes). Highly prevalent seed storage protein cDNAs were reduced (subtracted from the total cDNAs) by screening with the corresponding cDNAs.

Hybridization probes for screening the borage cDNA library were generated by using random primed DNA synthesis as described by Ausubel et al (1994 Current Protocols in Molecular Biology, Wiley Interscience, N.Y.) and corresponded to previously identified abundantly expressed seed storage protein cDNAs. Unincorporated nucleotides were removed by use of a G-50 spin column (Boehringer Manheim). Probe was denatured for hybridization by boiling in a water bath for 5 minutes, then quickly cooled on ice.

Nitrocellulose filters carrying fixed recombinant bacteriophage were prehybridized at 60°C for 2-4 hours in hybridization solution [4X SET (600 mM NaCl, 80 mM Tris-HCl, 4 mM Na<sub>2</sub>EDTA; pH 7.8), 5X Denhardt's reagent (0.1% bovine serum albumin, 0.1% Ficoll, and 0.1% polyvinylpyrolidone), 100 µg/ml denatured salmon sperm DNA, 50 µg/ml polyadenine and 10 ug/ml polycytidine]. This was replaced with fresh hybridization solution to which denatured radioactive probe (2 ng/ml

hybridization solution) was added. The filters were incubated at 60°C with agitation overnight. Filters

-25-

were washed sequentially in 4X, 2X, and 1X SET (150 mM NaCl, 20 mM Tris-HCl, 1 mM Na<sub>2</sub>EDTA; pH7.8) for 15 minutes each at 60°C. Filters were air dried and then exposed to X-ray film for 24 hours with intensifying screens at -80°C.

Non-hybridizing plaques were excised using Stratagene's excision protocol and reagents.

Resulting bacterial colonies were used to inoculate liquid cultures and were either sequenced manually or by an ABI automated sequencer.

10

# Random Sequencing of cDNAs from a Borage Seed 12 (DPP) Membrane-Bound Polysomal Library

Each cDNA corresponding to a nonhybridizing plaque was sequenced once and a sequence
tag generated from 200-300 base pairs. All sequencing
was performed by cycle sequencing (Epicentre). Over
300 expressed sequence tags (ESTs) were generated.
Each sequence tag was compared to GenBank database
using the BLAST algorithm (Altschul et al. 1990 J.

Mol. Biol. 215:403-410). A number of lipid metabolism
genes, including the Δ6-desaturase were identified.

Database searches with the cDNA clone designated mbp-65 using BLASTX with the GenBank database resulted in a significant match to the previously isolated Synechocystis A6-desaturase. It was determined however, that mbp-65 was not a full length cDNA. A full length cDNA was isolated using mbp-65 to screen the borage membrane-bound polysomal library. The resultant clone was designated pAN1 and the cDNA insert of pAN1 was sequenced by the cycle

-26-

sequencing method. The amino acid sequence deduced from the open reading frame (Fig. 1, SEQ ID NO:1) was compared to other known desaturases using Geneworks (IntelligGenetics) protein alignment program. This alignment indicated that the cDNA insert of pAN1 was the borage  $\Delta 6$ -desaturase gene.

The resulting dendrogram (Figure 2) shows that  $\Delta^{15}$ -desaturases and  $\Delta^{12}$ -desaturases comprise two groups. The newly isolated borage sequence and the previously isolated *Synechocystis*  $\Delta^{6}$ -desaturase (U.S.

- Patent No. 5,552,306) formed a third distinct group.
  A comparison of amino acid motifs common to
  desaturases and thought to be involved catalytically
  in metal binding illustrates the overall similarity of
  the protein encoded by the borage gene to desaturases
- in general and the *Synechocystis* Δ<sup>6</sup>-desaturase in particular (Table 1). At the same time, comparison of the motifs in Table 1 indicates definite differences between this protein and other plant desaturases. Furthermore, the borage sequence is also distinguished
- from known plant membrane associated fatty acid desaturases by the presence of a heme binding motif conserved in cytochrome b<sub>5</sub> proteins (Schmidt et al. 1994 Plant Mol. Biol. 26:631-642) (Figure 1). Thus, while these results clearly suggested that the
- isolated cDNA was a borage  $\Delta^6$ -desaturase gene, further confirmation was necessary. To confirm the identity of the borage  $\Delta 6$ -desaturase cDNA, the cDNA insert from pAN1 was cloned into an expression cassette for stable expression. The vector pBI121 (Jefferson et al. 1987)
- 30 EMBO J.  $\underline{6}$ :3901-3907) was prepared for ligation by

-27-

digestion with BamHI and EcoICR I (an isoschizomer of SacI which leaves blunt ends; available from Promega) which excises the GUS coding region leaving the 35S promoter and NOS terminator intact. The borage  $\Delta^6$ -desaturase cDNA was excised from the recombinant

plasmid (pAN1) by digestion with BamHI and XhoI. The XhoI end was made blunt by performing a fill-in reaction catalyzed by the Klenow fragment of DNA polymerase I. This fragment was then cloned into the BamHI/EcoICR I sites of pBI121.1, resulting in the

10 plasmid pAN2.

15

20

25

1		10x 2	FQIEHH (SEQ. ID. NO:17)	<b>Н</b> ОУТНН (SEQ. ID. NO:18)	HVIHH (SEQ. ID. NO:19)	HVIHH (SEQ. ID. NO:19)	(SEQ. ID. NO:19)	ID. NO:19)	ID. NO:19)	ID. NO:20)	ID. NO:20)	). NO:21)	ID. NO:21)	ID. NO:21)	ID. NO:21)	ID. NO:22)
	ASES	Metal Box 2	SEQ.	SEQ.	EQ. II	EQ. II	EQ. II	(SEQ. II	(SEQ. II	(SEQ. II	(SEQ. II	EQ. II	(SEQ. II	(SEQ. II	(SEQ. II	(SEQ. II
5	DESATURASES	Ħ	<b>Р</b> ОІВНН (	ннтуон (	в) нитин	нитин (s	нутнн (s	нутнн (s	нутнн (s	HVAHH (S	HVAHH (S	нірни (ѕеў. ір.	нтрин (s	з) ннатн	HIPHH (S	нирин (s
10	TE - BOUND	K 1	HNAHH (SEQ. ID. NO:9)	HNYLHH (SEQ. ID. NO:10)	HRTHH (SEQ. ID. NO:11)	NO:11)	NO:11)	NO:11)	NO:11)	NO:12)	NO:12)	NO:13)	NO:13)	NO:14)	NO:15)	NO:16)
	3RAN	Bos	ij	H	H.	ij.	Ë.	ID.	ij.	ij.	Ħ.	ij.	ïD.	ij.	ë.	9
	MEM	Metal Box 1	(SEQ.	SEQ	(SEQ.	(SEQ.	(SEQ.	(SEQ.	(SEQ.	(SEQ.	(SEQ.	(SEQ.	(SEQ.	(SEQ.	(SEQ.	(SEQ.
15	TABLE 1 MOTIFS IN		HNAHH	HNYLHH	нктин	HRTHH (SEQ. ID.	нктин (ѕер.	HRTHH	нктин (ѕео.	HRRHH	HRRHH	HDRHH (SEQ.	нркнн (ѕер.	нронн	нрнин	HNHHH
	TAI ACID M	Ų.	NO:3)	NO:4)	NO:5)	NO:5)	NO:5)	NO:5)	NO:5)	NO:6)	NO:6)	(7:ON	(7:ON	NO:7)	NO:8)	ID. NO:5)
20	ONI	Lipid Box	. B	. ib.	. H	. ID.	H.	H	i. ib.	. ID.	. ID.	i. ID.	. ID.	. ID.	. ID.	ë :
	N A	Lipi	ČES)	(SEQ.	(SEQ	(SEQ.	(SEQ	(SEQ	(SEQ	(SEQ	(SEQ.	(SEQ	(SEQ	(SEQ	(SEQ	(SEQ
	COMMO		WIGHDAGH (SEQ. ID. NO:3)	NVGHDANH	уденьсен (ѕер.	VLGHDCGH	VLGHDCGH (SEQ.	VLGHDCGH (SEQ.	уганосан (зер.	VIAHECGH (SEQ.	VIAHECGH	VIGHDCAH (SEQ.	VIGHDCAH (SEQ.	VIGHDCAH (SEQ.	VVGHDCGH (SEQ.	уъсноссн (ѕво.
25	N OF		æ	Z	5	5			Þ	>	>	Þ	Þ	>	5	5
30	TABLE 1 COMPARISON OF COMMON AMINO ACID MOTIFS IN MEMBRANE-BOUND	Desaturase	Borage A	Sweethorvatia A	Arah chloronlast A <sup>15</sup>	Rice A <sup>15</sup>	Glycine chloroplast A15	Arab, fad3 (A <sup>15</sup> )	Brassica fad 3 (A <sup>15</sup>	Borage A <sup>12</sup> (P1-81)*	Arab. fad2 (A <sup>12</sup> )	Arab, chloroplast A <sup>12</sup>	Clycine wleetid A <sup>12</sup>	Chinach plactidial n-6	Superhorvetia A <sup>12</sup>	Anabaena $\Delta^{12}$

\*P1-81 is a full length cDNA which was identified by EST analysis and shows high similarity to the Arbidopsis  $\Delta 12$  desaturase (fad2)

-29-

#### EXAMPLE 3

1

Production of Transgenic Plants and Preparation and Analysis of Fatty Acid Methyl Esters (FAMEs)

The expression plasmid, pAN2 was used to transform tobacco (Nicotiana tabacum cv. xanthi) via Agrobacterium tumefaciens according to standard procedures (Horsch, et al. 1985 Science 227:1229-1231; Bogue et al. 1990 Mol. Gen. Genet. 221:49-57) except that the initial transformants were selected on 100 μg/ml kanamycin.

Tissue from transgenic plants was frozen in liquid nitrogen and lyophilized overnight. FAMEs were prepared as described by Dahmer, et al. (1989) J.

15 Amer. Oil. Chem. Soc. 66: 543-548. In some cases, the solvent was evaporated again, and the FAMEs were resuspended in ethyl acetate and extracted once with deionized water to remove any water soluble contaminants. FAMEs were analyzed using a Tracor-560 gas liquid chromatograph as previously described (Reddy et al. 1996 Nature Biotech. 14:639-642).

As shown in Figure. 3, transgenic tobacco leaves containing the borage cDNA produced both GLA and octadecatetraenoic acid (OTA)  $(18:4 \Delta 6,9,12,15)$ .

25 These results thus demonstrate that the isolated cDNA encodes a borage  $\Delta 6$ -desaturase.

-30-

#### EXAMPLE 4

1 Expression of Δ6-desaturase in Borage

The native expression of A6-desaturase was examined by Northern Analysis of RNA derived from 5 borage tissues. RNA was isolated from developing borage embryos following the method of Chang et al. 1993 Plant Mol. Biol. Rep. 11:113-116. RNA was electrophoretically separated on formaldehydeagarose gels, blotted to nylon membranes by capillary  $^{10}$  transfer, and immobilized by baking at  $80^{\circ}\text{C}$  for  $^{30}$ minutes following standard protocols (Brown T., 1996 in Current Protocols in Molecular Biology, eds. Auselbel, et al. [Greene Publishing and Wiley-Interscience, New York] pp. 4.9.1-4.9.14.). The 15 filters were preincubated at 42°C in a solution containing 50% deionized formamide, 5X Denhardt's reagent, 5X SSPE (900 mM NaCl; 50mM Sodium phosphate, pH7.7; and 5 mM EDTA), 0.1% SDS, and 200  $\mu$ g/ml denatured salmon sperm DNA. After two hours, the  $^{20}$  filters were added to a fresh solution of the same composition with the addition of denatured radioactive hybridization probe. In this instance, the probes used were borage legumin cDNA (Fig. 16A), borage oleosin cDNA (Fig. 16B), and borage Δ6-desaturase cDNA  $^{25}$  (pAN1, Example 2) (Fig. 16C). The borage legumin and oleosin cDNAs were isolated by EST cloning and identified by comparison to the GenBank database using the BLAST algorithm as described in Example 2. Loading variation was corrected by normalizing to 30 levels of borage EF1 $\alpha$  mRNA. EF1 $\alpha$  mRNA was identified

-31-

by correlating to the corresponding cDNA obtained by the EST analysis described in Example 2. The filters were hybridized at 42°C for 12-20 hours, then washed as described above (except that the temperature was 65°C), air dried, and exposed to X-ray film.

As depicted in Figs. 15A and 15B, Δ6-desaturase is expressed primarily in borage seed. Borage seeds reach maturation between 18-20 days post pollination (dpp). Δ6-desaturase mRNA expression occurs throughout the time points collected (8-20 dpp), but appears maximal from 10-16 days post pollination. This expression profile is similar to that seen for borage oleosin and 12S seed storage protein mRNAs (Figs. 16A, 16B, and 16C).

15

5

20

25

-32-

### EXAMPLE 5

1 ISOLATION OF A SUNFLOWER ALBUMIN CDNA

The sunflower albumin cDNA (Ha5) was 5 isolated by differentially screening a sunflower cDNA library using cDNA probes from leaf and 12 DPF (days post flowering) embryos. A cDNA of 1011 bp was obtained (Cohen E.A. "Analysis of sunflower 2S seed storage protein genes" MS thesis, Texas A&M 10 University, Allen et al., 1987a in Molecular Approaches to Developmental Biology, pp. 415-424.). Although not full length, the cDNA comprised most of the coding sequence for the sunflower 2S albumin. Northern and dot blot analysis indicated that this 15 gene is exclusively expressed in developing sunflower seeds. Albumin transcripts and protein are first detected 5 DPF, a full two days earlier than helianthinin (11S), and reach maximal prevalence around 12-15 DPF (Allen et al. 1987a).

25

20

-33-

#### EXAMPLE 6

# ISOLATION OF A SUNFLOWER ALBUMIN 5' REGULATORY REGION

Genomic clones were isolated by screening a sunflower genomic DNA library using the Ha5 cDNA as a probe. Four independent genomic clones were shown to be identical by restriction enzyme digestion.

Therefore, one clone (HaG5) was chosen for more detailed analysis.

- A 2.3kb EcoRI/DraI fragment was sequenced

  (Allen et al., 1987b Mol. Gen. Genet. 210: 211-218).

  The HaG5 albumin gene contains two exons. The first exon (exon 1) is 575 nucleotides in length and the second exon (exon 2) is 310 nucleotides in length. A 190 nucleotide intron separates the two exons.
- Nuclease protection experiments showed that the transcription start site was located 30 nucleotides upstream of the translational start site. (Allen et al 1987b, Fig. 2). Southern analysis of genomic DNA and the fact that only one gene was isolated in an exhaustive screen indicated that HaG5 is a single copy gene in the sunflower genome.

An 889 bp upstream regulatory region (-860 to +29 of Figure 4; SEQ ID NO:2) was cloned in several steps from HaG5. A 1.1 kb EcoRI fragment was

25 subcloned in PBluescript™ (Stratagene) yielding pHaG5RI. PCR was performed on pHaG5RI with primers that resulted in the albumin 5' regulatory region being flanked by EcoRI and BamHI sites at the 5' and 3' ends, respectively. The restriction fragment was

30 cloned into the EcoRI/BamHI sites of pBluescript™

-34-

yielding pHaG5EB. Individual clones were sequenced to check possible PCR mutations as well as the orientation of their inserts. The sequence of the albumin 5' regulatory region is shown in Fig. 4 (SEQ. ID NO:2). The Sall/BamHI fragment of this construct was excised and cloned into pAN3 (the parental borage Δ6-desaturase containing plasmid), yielding pAN4. A map of pAN4 and the intermediate vectors involved in its construction are shown in Fig. 5. pAN1 is described in Example 2. pBI101.1 is described in (Jefferson et al. 1987 EMBO J. 6:3901-3907).

-35-

## EXAMPLE 7

EXPRESSION OF \( \triangle 6 - \text{DESATURASE UNDER CONTROL OF THE SUNFLOWER ALBUMIN 5' REGULATORY REGION \)

The albumin 5' regulatory region was used to drive the expression of a borage  $\Delta 6$ -desaturase gene in Arabidopis. pAN4 was used to transform Arabidopsis using the vacuum infiltration method of Bechtold et al. 1993 C.R. Acad. Sci. Paris 316: 1194-1199. Levels of  $\Delta 6$ -desaturase activity were monitored by assaying the corresponding fatty acid methyl esters of its reaction products, y-linolenic acid (GLA) and octadecatetraenoic acid (OTA) using the methods described in Example 3. GLA and OTA levels in transgenic seeds ranged up to 10.2% (average of 4.4%) and 3.6% (average of 1.7%), respectively, of the C18 15 fatty acids. No GLA or OTA was detected in the leaves of these plants. In comparison, 35S promoter/ $\Delta^6$ desaturase transgenic plants produced GLA levels of up to 3.1% of C18 fatty acids (average of 1.3%) in leaves and no measurable OTA in seeds. These data are 20 summarized in Table 2.

25

-36-

1

5

10

15

TABLE 2 EXPRESSION OF THE BORAGE  $\Delta^6\text{-} \text{DESATURASE}$  IN TRANSGENIC PLANTS

20

25

OTA 7.6 n.d LEAF RANGE 19-22 GLA RANGE 20 8-11 n.d. OTA\* n.d SEED 3.1-10.2 0.7-3.1 RANGE GLA\* RANGE 4.4 0.63-1.3 Arabidops is PLANT tobacco Cauliflower mosaic virus 35S Sunflower albumin PROMOTER

\*mean value expressed as the percent of the  $\mathsf{C}_{18}$  fatty acids n.d. not detected

-37-

#### EXAMPLE 8

Transformation of Oilseed Rape With an Expression Cassette Which Comprises the Albumin 5' Regulatory Region Linked to the Borage Delta 6-Desaturase Gene

Oilseed rape, Cv. Westar, was transformed with the strain of Agrobacterium tumefaciens EHA105 containing the plasmid pAN4 (i.e. the borage  $\Delta 6$ -desaturase gene under the control of the sunflower albumin promoter-Example 6).

cultivated for 2-3 days with induced Agrobacterium tumefaciens strain EHA105 (Alt-Moerbe et al. 1988 Mol. Gen. Genet. 213:1-8; James et al. 1993 Plant Cell Reports 12:559-563), then transferred onto regeneration medium (Boulter et al. 1990 Plant Science 70:91-99; Fry et al. 1987 Plant Cell Reports 6:321-325). The regenerated shoots were transferred to growth medium (Pelletier et al. 1983 Mol.Gen. Menet. 191:244-250), and a polymerase chain reaction (PCR) test was performed on leaf fragments to assess the presence of the gene.

DNA was isolated from the leaves according to the protocol of K.M. Haymes et al. (1996) Plant Molecular Biology Reporter, 14(3):280-284, and resuspended in 100µl of water, without RNase 25 treatment. 5µl of extract were used for the PCR reaction, in a final volume of 50µl. The reaction was performed in a Perkin-Elmer 9600 thermocycler, with the following cycles:

-38-

1 cycle: 95°C, 5 minutes

1 30 cycles: 95°C, 45 sec; 52°C, 45 sec

72°C, 1 minute

1 cycle: 72°C, 5 minutes

5 and the following primers (derived from near the metal box regions, as indicated in Fig. 1, SEQ. NO.:1):

5' TGG AAA TGG AAC CAT AA 3'

5' GGA AAC AAA TGA TGC TC 3'

Amplification of the DNA revealed the expected 549

10 base pair PCR fragment (Figure 7).

The positive shoots were transferred to elongation medium, then to rooting medium (DeBlock et al 1989 *Plant Physiol.* 91:694-701). Shoots with a well-developed root system were transferred to the

greenhouse. When plants were well developed, leaves were collected for Southern analysis in order to assess gene copy number.

Genomic DNA was extracted according to the procedure of Bouchez et al. (1996) Plant Molecular

Biology Reporter 14:115-123, digested with the restriction enzymes Bgl I and/or Cla I, electrophoretically separated on agarose gel (Maniatis et al. 1982, in Molecular Cloning; a Laboratory manual. Cold Spring Harbor Laboratory Press, Cold

- 25 Spring Harbor/NY), and prepared for transfer to nylon membranes (Nytran membrane, Schleicher & Schuell) according to the instructions of the manufacturer.

  DNA was then transferred to membranes overnight by capillary action using 20XSSC (Maniatis et al. 1982).
- 30 Following transfer, the membranes were crosslinked by

-39-

UV (Stratagene) for 30 seconds and pre-hybridized for 1 hour at 65°C in 15 ml of a solution containing 6xssc, 0.5%sDs and 2.25% w/w dehydrated skim milk in glass vials in hybridization oven (Appligene). The membranes were hybridized overnight in the same

- 5 solution containing a denatured hybridization probe radiolabelled with <sup>32</sup>P to a specific activity of 10<sup>8</sup> cpm/µg by the random primer method (with the Ready-To-Go kit obtained from Pharmacia). The probe represents a PCR fragment of the borage delta 6-desaturase gene
- 10 (obtained in the conditions and with the primers detailed above). After hybridization, the filters were washed at 65°C in 2XSSC, 0.1% SDS for 15 minutes, and 0.2XSSC, 0.1%SDS for 15 minutes. The membranes were then wrapped in Saran-Wrap and exposed to Kodak
- 15 XAR film using an intensifying screen at -70°C in a, light-proof cassette. Exposure time is generally 3 days.

The results obtained confirm the presence of the gene. According to the gene construct, the number of bands in each lane of DNA digested by Bgl I or Cla I represents the number of delta 6-desaturase genes present in the genomic DNA of the plant. The digestion with Bgl I and Cla I together generates a fragment of 3058 bp.

- The term "comprises" or "comprising" is defined as specifying the presence of the stated features, integers, steps, or components as referred to in the claims, but does not preclude the presence or addition of one or more other features, integers, steps,
- 30 components, or groups thereof.

WO 98/45460

PCT/US98/07178

-40-

## SEOUENCE LISTING

#### (1) GENERAL INFORMATION:

- (i) APPLICANT: Rhone Poulenc Agro
  Thomas, Terry L.
  Nunberg, Andrew N.
  Beremand, Phillip D.
- (ii) TITLE OF INVENTION: A SUNFLOWER ALBUMIN 5' REGULATORY REGION FOR THE MODIFICATION OF PLANT SEED LIPID COMPOSITION
- (iii) NUMBER OF SEQUENCES: 22
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Scully, Scott, Murphy & Presser
  - (B) STREET: 400 Garden City Plaza
  - (C) CITY: Garden City
  - (D) STATE: New York
  - (E) COUNTRY: USA
  - (F) ZIP: 11530
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: 08/831,570
  - (B) FILING DATE: 09-APR-1997
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: DiGiglio, Frank S.
  - (B) REGISTRATION NUMBER: 31,346
  - (C) REFERENCE/DOCKET NUMBER: 10545
  - (ix) TELECOMMUNICATION INFORMATION:
    - (A) TELEPHONE: (516) 742-4343
    - (B) TELEFAX: (516) 743-4366

-41-

(2) INFORMATION	FOR	SEQ	ID	NO:1:	:
-----------------	-----	-----	----	-------	---

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1684 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double (D) TOPOLOGY: linear

# (ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 43..1387

# (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATAT	rctg(	CCT A	ACCCI	rccc <i>i</i>	AA AC	GAGA	STAGT	CA7	r <b>TT</b> T	CAT		ATG ( Met <i>I</i> 1				54
ATC Ile 5	AAG Lys	AAA Lys	TAC Tyr	ATT Ile	ACC Thr 10	TCA Ser	GAT Asp	GAA Glu	CTC Leu	AAG Lys 15	AAC Asn	CAC His	GAT Asp	AAA Lys	CCC Pro 20	102
GGA Gly	GAT Asp	CTA Leu	TGG Trp	ATC Ile 25	TCG Ser	ATT Ile	CAA Gln	GGG Gly	AAA Lys 30	GCC Ala	TAT Tyr	GAT Asp	GTT Val	TCG Ser 35	GAT Asp	150
TGG Trp	GTG Val	AAA Lys	GAC Asp 40	CAT His	CCA Pro	GGT Gly	GGC Gly	AGC Ser 45	TTT Phe	CCC Pro	TTG Leu	AAG Lys	AGT Ser 50	CTT Leu	GCT Ala	198
GGT Gly	CAA Gln	GAG Glu 55	GTA Val	ACT Thr	GAT Asp	GCA Ala	TTT Phe 60	GTT Val	GCA Ala	TTC Phe	CAT His	CCT Pro 65	GCC Ala	TCT Ser	ACA Thr	246
TGG Trp	AAG Lys 70	AAT Asn	CTT Leu	GAT Asp	AAG Lys	TTT Phe 75	TTC Phe	ACT Thr	GGG Gly	TAT Tyr	TAT Tyr 80	CTT Leu	AAA Lys	GAT Asp	TAC Tyr	294
TCT Ser 85	GTT Val	TCT Ser	GAG Glu	GTT Val	TCT Ser 90	AAA Lys	GAT Asp	TAT Tyr	AGG Arg	AAG Lys 95	CTT Leu	GTG Val	TTT Phe	GAG Glu	TTT Phe 100	342
TCT Ser	AAA Lys	ATG Met	GGT Gly	TTG Leu 105	TAT Tyr	GAC Asp	AAA Lys	AAA Lys	GGT Gly 110	CAT His	ATT Ile	ATG Met	TTT Phe	GCA Ala 115	ACT Thr	390

-42-

TTG Leu	TGC Cys	TTT Phe	ATA Ile 120	GCA Ala	ATG Met	CTG Leu	TTT Phe	GCT Ala 125	ATG Met	AGT Ser	GTT Val	TAT Tyr	GGG Gly 130	GTT Val	TTG Leu		438
TTT Phe	TGT Cys	GAG Glu 135	GGT Gly	GTT Val	TTG Leu	GTA Val	CAT His 140	TTG Leu	TTT Phe	TCT Ser	GGG Gly	TGT Cys 145	TTG Leu	ATG Met	GGG Gly		486
TTT Phe	CTT Leu 150	TGG Trp	ATT Ile	CAG Gln	AGT Ser	GGT Gly 155	TGG Trp	ATT Ile	GGA Gly	CAT His	GAT Asp 160	GCT Ala	GGG Gly	CAT His	TAT Tyr		534
ATG Met 165	GTA Val	GTG Val	TCT Ser	GAT Asp	TCA Ser 170	AGG Arg	CTT Leu	AAT Asn	AAG Lys	TTT Phe 175	ATG Met	GGT Gly	ATT Ile	TTT Phe	GCT Ala 180		582
GCA Ala	AAT Asn	TGT Cys	CTT Leu	TCA Ser 185	GGA Gly	ATA Ile	AGT Ser	ATT Ile	GGT Gly 190	TGG Trp	TGG Trp	AAA Lys	TGG Trp	AAC Asn 195	CAT His		630
AAT Asn	GCA Ala	CAT His	CAC His 200	ATT Ile	GCC Ala	TGT Cys	AAT Asn	AGC Ser 205	CTT Leu	GAA Glu	TAT Tyr	GAC Asp	CCT Pro 210	GAT Asp	TTA Leu		678
CAA Gln	TAT Tyr	ATA Ile 215	CCA Pro	TTC Phe	CTT Leu	GTT Val	GTG Val 220	TCT Ser	TCC Ser	AAG Lys	TTT Phe	TTT Phe 225	GGT Gly	TCA Ser	CTC Leu		726
ACC Thr	TCT Ser 230	CAT His	TTC Phe	TAT Tyr	GAG Glu	AAA Lys 235	AGG Arg	TTG Leu	ACT Thr	TTT Phe	GAC Asp 240	TCT Ser	TTA Leu	TCA Ser	AGA Arg		774
TTC Phe 245	TTT Phe	GTA Val	AGT Ser	TAT Tyr	CAA Gln 250	CAT His	TGG Trp	ACA Thr	TTT Phe	TAC Tyr 255	CCT Pro	ATT Ile	ATG Met	TGT Cys	GCT Ala 260		822
GCT Ala	AGG Arg	CTC Leu	AAT Asn	ATG Met 265	TAT Tyr	GTA Val	CAA Gln	TCT Ser	CTC Leu 270	ATA Ile	ATG Met	TTG Leu	TTG Leu	ACC Thr 275	AAG Lys		870
AGA Arg	AAT Asn	GTG Val	TCC Ser 280	TAT Tyr	CGA Arg	GCT Ala	CAG Gln	GAA Glu 285	CTC Leu	TTG Leu	GGA Gly	TGC Cys	CTA Leu 290	GTG Val	TTC Phe		918
TCG Ser	ATT Ile	TGG Trp 295	TAC Tyr	CCG Pro	TTG Leu	CTT Leu	GTT Val 300	TCT Ser	TGT Cys	TTG Leu	CCT Pro	AAT Asn 305	TGG Trp	GGT Gly	GAA Glu		966
AGA Arg	ATT Ile 310	ATG Met	TTT Phe	GTT Val	ATT Ile	GCA Ala 315	AGT Ser	TTA Leu	TCA Ser	GTG Val	ACT Thr 320	GGA Gly	ATG Met	CAA Gln	CAA Gln	1	1014

-43-

			TCC Ser													1062	?
			AAT Asn													1110	)
			CCT Pro 360													1158	3
CAA Gln	ATT Ile	GAG Glu 375	CAT His	CAT His	TTG Leu	TTT Phe	CCC Pro 380	AAG Lys	ATG Met	CCT Pro	AGA Arg	TGC Cys 385	AAC Asn	CTT Leu	AGG Arg	1206	5
AAA Lys	ATC Ile 390	TCG Ser	CCC Pro	TAC Tyr	GTG Val	ATC Ile 395	GAG Glu	TTA Leu	TGĆ Cys	AAG Lys	AAA Lys 400	CAT His	AAT Asn	TTG Leu	CCT Pro	1254	1
TAC Tyr 405	AAT Asn	TAT Tyr	GCA Ala	TCT Ser	TTC Phe 410	TCC Ser	AAG Lys	GCC Ala	AAT Asn	GAA Glu 415	ATG Met	ACA Thr	CTC Leu	AGA Arg	ACA Thr 420	1302	2
			ACA Th'r													1350	)
			GTA Val 440										АААТ'	TACC	2	1397	7
TTA	GTTC	ATG '	TAAT	AATT'	rg A	GATT	ATGT	A TC	TCCT:	ATGT	TTG	TGTC'	TTG '	rctt(	GGTTCT	145	7
ACT!	rgtt	GGA (	GTCA'	TTGC	AA C'	rtgt(	CTTT	r AT	GGTT'	TATT	AGA'	TGTT'	TTT '	TAAT	ATATTT	151	7
TAG	AGGT'	rtt (	GCTT'	rcat(	CT C	CATT	ATTG	A TG	AATA	AGGA	GTT	GCAT.	ATT	GTCA	ATTGTT	157	7
GTG	CTCA	ATA	TCTG	АТАТ'	TT T	GGAA'	TGTA(	C TT	TGTA	CCAC	GTG	GTTT	TCA	GTTG:	AAGCTC	163	7
ATG'	rgta(	CTT ·	CTAT	AGAC'	TT T	GTTT.	TAAA	g gt	TATG'	TCAT	GTT.	ATTT				168	4

# (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 843 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double(D) TOPOLOGY: linear

-44-

(ii)	MOLECULE	TYPE:	DNA	(genomic)
------	----------	-------	-----	-----------

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2: GAATTCTATC ACTAGTGACC ACCCCATCCC CTTATTTCAA TAATGGAACA CAAAAAATT 60 TTAAAAAATA GTTGCTGTTA ATTGTTTAAC CGTCATTTTC CAACATTACT AGCTAATCGT 120 TAATTGATCT TCATAAAAA AAAAATTGCT ATGGGTACTA TTGAGATTGT ATATCTTATC 180 AGTTAGGCCT AAGGGGGGGG TCAGTGATAT TACGAATGAT ACAAACATCA ACGCGTGGAA 240 300 TGATGGTAAT TGTTGGTTGG GGGGAAATTA TTGGGTATGG TGTTGAGTAT GACCACCCC 360 ACTAAAAAAG GTTGTGAGTG ATGTAAAAAT GGTTGCTGAC ATGACGAAAC ATAATTGGAT 420 ATTGTGAGTG ATAAAATTTT ATCATTAGTG ACCACCCCGC CTCCCCTTAT CATATGTTGT 480 TATCTTCCAT AGTTGCGGTA TACCAACATA TGGTAGTTTT TATATTTATA GTTTATATTT 540 TCATTAAACT CTCTTCGCCA GGCTACTTGT ATTGTAATCA TATGGAATCT CAACTCCAGT 600 TGGAGCCATT CCATCATATA TTTCCATTTC CAAACAAAGA GAATTGACAC CTCATACATA 660 CTCCAAAGCA TACTTCCACT TGCTATAATT TTCATGTAAA AACTCGTACG TGTTATTCGA 720 CAATGTTCAT ATAACGCCAC CGATTAAACT CACCTCTCCA CGTATGAACC TCCACCCACC 780 ATATATACGC ACCACCACCA CACCATAATT CACACAACCA CAACACCATC TCCCACAGGA 840

843

## (2) INFORMATION FOR SEQ ID NO:3:

TCC

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Trp Ile Gly His Asp Ala Gly His 1

WO 98/45460

-45-

- (2) INFORMATION FOR SEQ ID NO:4:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 8 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Asn Val Gly His Asp Ala Asn His 1

- (2) INFORMATION FOR SEQ ID NO:5:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 8 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Val Leu Gly His Asp Cys Gly His 1 5

- (2) INFORMATION FOR SEQ ID NO:6:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 8 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Val Ile Ala His Glu Cys Gly His 1 5 -46-

- (2) INFORMATION FOR SEQ ID NO:7:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 8 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Val Ile Gly His Asp Cys Ala His 1 5

- (2) INFORMATION FOR SEQ ID NO:8:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 8 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Val Val Gly His Asp Cys Gly His 1

- (2) INFORMATION FOR SEQ ID NO:9:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 5 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

His Asn Ala His His 1 5

-47-

- (2) INFORMATION FOR SEQ ID NO:10:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 6 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

His Asn Tyr Leu His His 1 5

- (2) INFORMATION FOR SEQ ID NO:11:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 5 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

His Arg Thr His His 1 5

- (2) INFORMATION FOR SEQ ID NO:12:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 5 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

His Arg Arg His His 1 5

WO 98/45460

-48-

- (2) INFORMATION FOR SEQ ID NO:13:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 5 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

His Asp Arg His His 1 5

- (2) INFORMATION FOR SEQ ID NO:14:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 5 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

His Asp Gln His His 1 5

- (2) INFORMATION FOR SEQ ID NO:15:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 5 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

His Asp His His His 1

-49-

- (2) INFORMATION FOR SEQ ID NO:16:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 5 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

His Asn His His His 1

- (2) INFORMATION FOR SEQ ID NO:17:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 6 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Phe Gln Ile Glu His His 1 5

- (2) INFORMATION FOR SEQ ID NO:18:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 6 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

His Gln Val Thr His His 1 5

- (2) INFORMATION FOR SEQ ID NO:19:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 6 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

His Gln Val Thr His His 1 5

- (2) INFORMATION FOR SEQ ID NO:20:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 5 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

His Val Ala His His 1 5

- (2) INFORMATION FOR SEQ ID NO:21:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 5 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

His Ile Pro His His 1 5

-51-

- (2) INFORMATION FOR SEQ ID NO:22:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 5 amino acids(B) TYPE: amino acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

His Val Pro His His

-52-

## What is claimed is:

l

- 1. An isolated nucleic acid encoding an albumin 5' regulatory region which directs seed-specific expression selected from the group consisting of the nucleotide sequence set forth in SEQ ID NO:2, the nucleotide sequence set forth in SEQ ID NO:2 having an insertion, deletion, or substitution of one or more nucleotides, and a contiguous fragment of the nucleotide sequence set forth in SEQ ID NO:2.
- 2. An expression cassette which comprises the albumin 5' regulatory region of Claim 1 operably linked to a heterologous gene.
- 3. The expression cassette of Claim 2 wherein the heterologous gene is at least one of a fatty acid synthesis gene or a lipid metabolism gene.
- 4. The expression cassette of Claim 3 wherein the heterologous gene is selected from the group consisting of a lipid desaturase gene, an acyl carrier protein (ACP) gene, a thioesterase gene, an acetyl transacylase gene, an acetyl-coA carboxylase gene, a ketoacyl synthase gene, a malonyl transacylase gene or an elongase gene.
- 5. The expression cassette of Claim 4 wherein the lipid desaturase gene is selected from the 25 group consisting of a  $\Delta 6$ -desaturase gene, a  $\Delta 12$ -desaturase gene, and a  $\Delta 15$ -desaturase gene.
  - 6. An expression vector which comprises the expression cassette of any one of Claims 2-5.
- 7. A cell comprising the expression 30 cassette of any one of Claims 2-5.

-53-

- 8. A cell comprising the expression vector  $\ensuremath{\text{l}}$  of Claim 6.
  - 9. The cell of Claim 7 wherein said cell is a bacterial cell or a plant cell.
- 10. The cell of Claim 8 wherein said cell 5 is a bacterial cell or a plant cell.
  - 11. A transgenic plant comprising the expression cassette of any one of Claims 2-5.
  - 12. A transgenic plant comprising the expression vector of Claim 6.
- 13. A plant which has been regenerated from the plant cell of Claim 9.
  - 14. A plant which has been regenerated from the plant cell of Claim 10.
- 15. The plant of Claim 12 or 13 wherein said plant is at least one of a sunflower, soybean, maize, cotton, tobacco, peanut, oil seed rape or Arabidopisis plant.
  - 16. Progeny of the plant of Claim 11 or 12.
  - 17. Seed from the plant of Claim 11 or 12.
- 20 18. A method of producing a plant with increased levels of a product of a lipid metabolism gene which comprises:
  - (a) transforming a plant cell with an expression vector comprising the isolated nucleic acid of Claim 1 operably linked to at least one of an
- of Claim 1 operably linked to at least one of an isolated nucleic acid coding for a fatty acid synthesis gene or a lipid metabolism gene; and
  - (b) regenerating a plant with increased levels of the product of said fatty acid synthesis

- 54 -

gene or said lipid metabolism gene from said plant l cell.

- 19. A method of producing a plant with increased levels of gamma linolenic acid (GLA) content which comprises:
- 5 (a) transforming a plant cell with an expression vector comprising the isolated nucleic acid of Claim 1 operably linked to a A6-desaturase gene; and
- (b) regenerating a plant with increased 10 levels of GLA from said plant cell.
  - 20. The method of Claim 19 wherein said  $\Delta 6$ desaturase gene is at least one of a cyanobacterial  $\Delta 6$ -desaturase gene or a borage  $\Delta 6$ -desaturase gene.
- 21. The method of Claim 18 or 19 wherein 15 said plant is a sunflower, soybean, maize, tobacco, cotton, peanut, oil seed rape or Arabidopsis plant.

22.

- The method of Claim 18 wherein said fatty acid synthesis gene or said lipid metabolism gene is at least one of a lipid desaturase, an acyl 20 carrier protein (ACP) gene, a thioesterase gene an elongase gene, an acetyl transacylase gene, an acetylcoA carboxylase gene, a ketoacyl synthase gene, or a malonyl transacylase gene.
- 23. A method of inducing production of at 25 least one of gamma linolenic acid (GLA) or octadecatetraeonic acid (OTA) in a plant deficient or lacking in GLA which comprises transforming said plant with an expression vector comprising an the isolated nucleic acid of Claim 1 operably linked to a  $\Delta 6$ -

-55-

desaturase gene and regenerating a plant with l increased levels of at least one of GLA or OTA.

- 24. A method of decreasing production of a fatty acid synthesis or lipid metabolism gene in a plant which comprises:
- (a) transforming a plant cell with an expression vector comprising the isolated nucleic acid of Claim 1 operably linked to a nucleic acid sequence complementary to a fatty acid synthesis or lipid metabolism gene; and
- 10 (b) regenerating a plant with decreased production of said fatty acid synthesis or said lipid metabolism gene.
- 25. A method of cosuppressing a native fatty acid synthesis or lipid metabolism gene in a plant which comprises:
- (a) transforming a cell of the plant with an expression vector comprising the isolated nucleic acid of Claim 1 operably linked to a nucleic acid sequence encoding a fatty acid synthesis or lipid metabolism
   20 gene native to the plant; and
  - (b) regenerating a plant with decreased production of said fatty acid synthesis or said lipid metabolism gene.

25

ara tot god tad oot ood naa gag agt agt cat tit toa toa alg got got caa ald aag 62 ass the att acc toe get ges ote sag sac can get ass one ggs get ots tog sto top K Y I T S D K L K H H D K P G D L W I S 122 coc ttg aag agt ctt gct ggt caa gag gta act gat gcs ttt gtt gcs ttc cat cct gcc P L K S L  $\lambda$  G Q E V T D  $\lambda$  F V  $\lambda$  F R P  $\lambda$ 9 T W K H I, D K F F T G Y Y I, K D Y S V 102 tot aca top mag mat oft gat mag tit tio act gog tat tat oft mam gat the tot git tet gag get tet asa gat tat agg aag ett geg tit gag tet tet aaa atg ggt tig tat S S V S K D Y R K L V F S F S K H G L Y 362 gac ass ass got cat att atg ttt gos act ttg tgg ttt ats gos atg ctg ttt got atg D K K G H I H F  $\lambda$  T  $\nu$  C F I  $\lambda$  H  $\nu$  F  $\lambda$  H 422 agt git tat gog git tig tit tot gag got git tig gia cat tig tit tet gog tgt tig S V Y G V L F C B G V L V II L F S G C L 482 atg ggg ttt ctt tgg att cag agt ggt tgg att gga cat gat gct ggg cat tat atg gta H G F L H I Q S G  $\frac{H}{H}$  J G  $\frac{H}{H}$  D A G  $\frac{H}{H}$  Y H V gtg tot get toe egg ctt eat mag ttt atg ggt att ttt gct gca aat tgt ctt tce gga V S D S R  $I_1$  N K P M G I F A A N C  $I_2$  S Gat another top top top and top and cat ast goalest can att gool tot and age of the set gam the gac cot gat the cam tat sta com the cet get get too mag tet tet ggt B Y D P D L Q Y I P F L V V S S K F F G tea etc acc tet cat the tat gag ass agg tig act tit gae tet tis tes ags tie tit S L T S H F Y E K R L T F D S L S R F F 782 gia agt tat cas cat tgg aca itt tac cet att atg tgt get get agg etc ant atg tat V  $\cdot$ S Y Q II H T F Y P I H C A A R I, N H Y 842 gia caa tet ete ats atg tig teg ace aag aga aat gig tee tat ega get cag gas ete V  $\mathbb Q$  S  $\mathbb C$   $\mathbb$ ggt gam agm att mig tit git att gca mig tim tom gig mit ggm mig cam cam git eng G R R I H P V I A S I, S V T G H Q Q V Q 1022 tto tee ttg sac can tto tet ten agt gtt tat gtt ggs sag eet amm ggg mat amt tgg F S L N N F S S g V Y V G K P K G N N W 1082 ttt gag amm cam ocg gat gog och ett gad att tot tot oct oct tog atg gat tog ttt EKQTDGTLDISCPPWH, cat got goa tig cam itc cam att gag cat cat itg itt ccc mag mig cct mag itg mac iii G G L O F Q I R II L F P K H P R C H 1202 ctt agg man atc tog occ the gtg ste gag ttn tgc mag ann cat mat ttg cot the ant L R K I S P Y V I S L C K K H H L P Y H 1262 tat gcs tot tot toe and gcc sat gas stg acs ote ags acs ttg agg sac acs gcs ttg Y Å S P S K  $\lambda$  N B M T L R T L R N T  $\lambda$  L 1322 end jet and gat at a nee say cod etc cod mag sat ttg gta tgg gas get ett car set Q  $\lambda$  R  $\nu$  T  $\nu$  F  $\nu$  F 1302 cat got tam amt tag got tag tto mtg tam tom ttt gag att atg tat etc eta tgt ttg 1442 tgt oft gio tig git ofa cit get gga gio att gca act tgt cit ita tgg tit att aga tgt tit tie ata thi tit aga ggt tit get tie nie tee att att gni gwa tha gga git 1562 goa tat tot cas tin the toe tes sta tot eat att the eas tot act the tac cae tot 1622 get the age tga age ten tgt gea ett eta tag act tig tit aan tgg tea tgt eat git

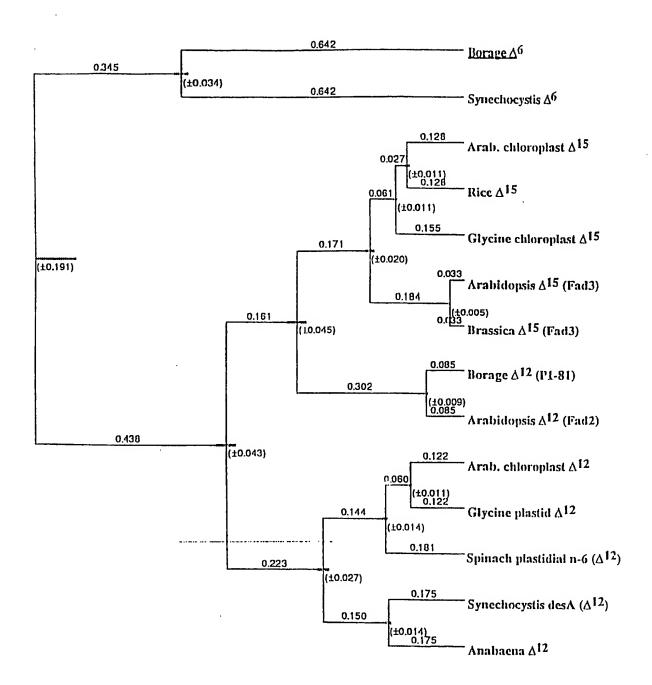


FIGURE 2

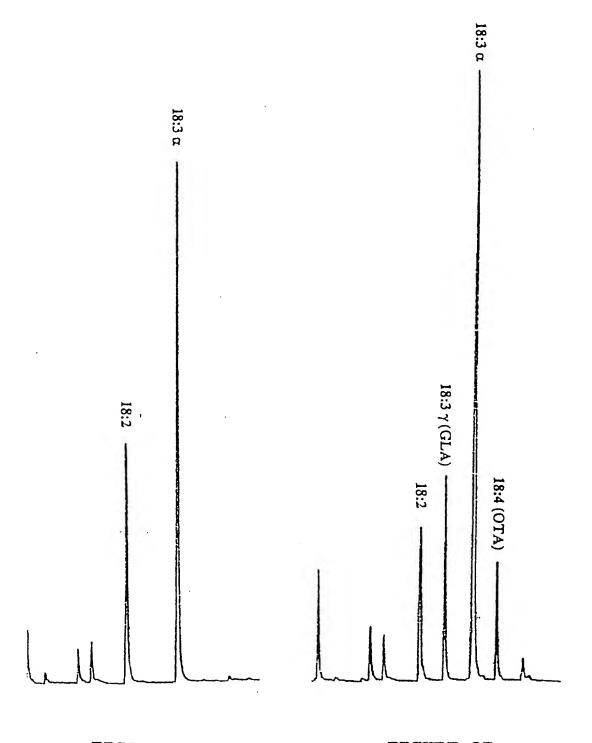
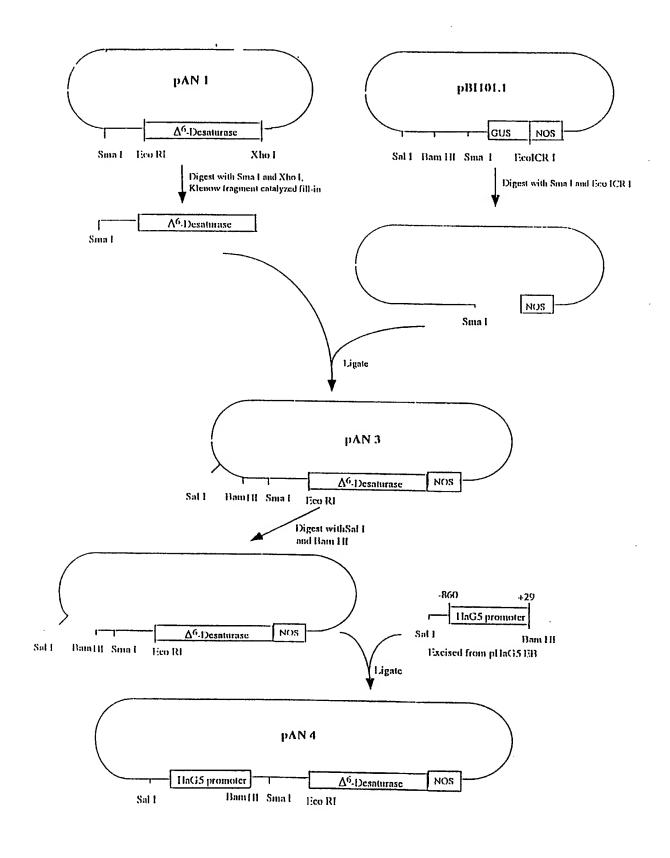


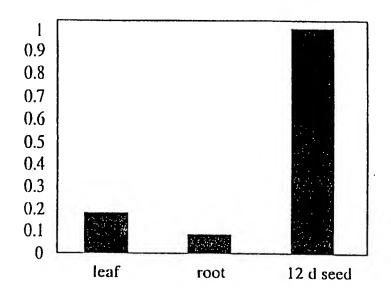
FIGURE 3A

FIGURE 3B

TGINIC VCINGIBYCC YCCCCNICCC GIINIIIICYY INYIGGYYCY	-811
AAAAAT' TITI'AAAAAAT' AGI'IGC'IGIT' AA'I'IGI'ITI'AA CCGICATI'ITI'	-761
CATTING TAGCTANTOG TTANTIGATO TICATANAAAA AAAAANTTOO	-711
OGINCT NITIGAGNITIG TATATCITAT CAGTTAGGCC TAAGGGGCCG	-661
GIGATA TIPACGAATGA TACAAACATC AACGCGIGGA ACATTACAAA	-611
TAILCCC CACCICCAAG TATAACGCGF GITTGFTCCA CGGFFIGATG	-561
CGTANT TTTTTCAACG CCGIGNTGGF TTTTTTTTTTT TTTTTTTTTT	-51.1
LICGIAN TIGITICGUIG COCCGANNTU NUTCCGUNIG GUGUTGAGUG	-461
CCACCC CCACLANNAN AGGLIGIGAG IGNIGINANA NIGGLIGCIG	-411
GACGAA ACATAATTOG ATATTGIGAG TGATAAAATT TTATCATTAG	-361
CACCCC GCCICCCCIT ATCATAIGIT GITATCLICC ATAGITGCGG	-311
CCNACA TATGGI'NGI'I' TTTATATI'AT TAGTI'I'ATAT TITCATI'AAA	-261
CI'ICCC CACCCI'ACTT GI'ATTGTAAT CATAIGGAAT CI'CAACI'CCA	-211
GAGCCA TICCNICAIN TATTICCATT TCCAAACAAA GAGANITGAC	-161
CNINCA TACICCAAAG CATACITCCA CITOCTATAA TITICAIGTA	-111
CICGIA CGIGITATIC GACAATGITC ATATAACGCC ACCGATTAAA	-61.
ICCICIC CYCCIAIGAY CCICCYCCCY CCAIMININIAC GCYCCYCCYC	-1.1
VCCN'I'NA 'I' I'CACACANC CACAACACCA 'I'CICCCACA <u>G GA'ICC</u>	+29



# FIGURE 6B

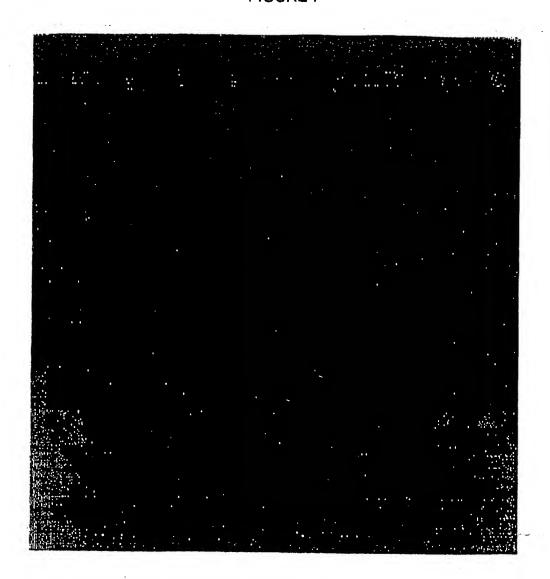


# Borage tissue



FIGURE 6A

FIGURE 7



Interny val Application No PCT/US 98/07178

A. CLASSIFICATION OF SUBJECT MATTER IPC 6 C12N15/82 C12N15/29 C12N15/53 A01H5/00 A01H5/10 According to International Patent Classification (IPC) or to both national classification and IPC **B. FIELDS SEARCHED** Minimum documentation searched (classification system followed by classification symbols) IPC 6 C12N A01H Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practical, search terms used) C. DOCUMENTS CONSIDERED TO BE RELEVANT Citation of document, with indication, where appropriate, of the relevant passages Relevant to claim No. Category \* ALLEN, R.D., ET AL.: "Sequence and 1 X expression of a gene encoding an albumin storage protein in sunflower' MOL. GEN. GENET., vol. 210, 1987, pages 211-218, XP002074856 2-23 see the whole document WO 96 21022 A (RHONE POULENC AGROCHIMIE) 2-23 11 July 1996 see page 14, line 3 - page 15, line 2; examples 6,11,13,14 WO 92 17580 A (RHONE POULENC AGROCHIMIE) 1-25 A 15 October 1992 see page 20, line 1 - page 11 WO 94 10189 A (CALGENE INC) 11 May 1994 1-4 see examples 3,4 -/--Further documents are listed in the continuation of box C. Х Patent family members are listed in annex. X Special categories of cited documents: "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the "A" document defining the general state of the art which is not considered to be of particular relevance invention "E" earlier document but published on or after the international "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such docu-"O" document referring to an oral disclosure, use, exhibition or ments, such combination being obvious to a person skilled in the art. document published prior to the international filing date but later than the priority date claimed "&" document member of the same patent family Date of the actual completion of the international search Date of mailing of the international search report 0 7, 09, 98 4 September 1998 **Authorized officer** Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentiaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Maddox. A Fax: (+31-70) 340-3016

Intern 1al Application No PCT/US 98/07178

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Quienna èn alaim Ma
	une instant passages	Relevant to claim No.
A	WO 91 13972 A (CALGENE INC) 19 September 1991 see examples 9,13	1-4
A	WO 92 11373 A (DU PONT) 9 July 1992 see page 67, line 25 - page 68, line 13	1-4
A	WO 93 11245 A (DU PONT) 10 June 1993 see examples 10-13	1-5,24
A	WO 94 18337 A (MONSANTO CO ;UNIV MICHIGAN (US); GIBSON SUSAN IRMA (US); KISHORE G) 18 August 1994 see the whole document	1,3-5, 24,25
A	WO 94 11516 A (DU PONT ;LIGHTNER JONATHAN EDWARD (US); OKULEY JOHN JOSEPH (US)) 26 May 1994 see page 98 - page 110	1,3-5,24
A	WO 96 06936 A (DU PONT ;HITZ WILLIAM DEAN (US)) 7 March 1996 see page 48, line 23 - page 53, line 10	1-5,24, 25
A	BEREMAND, P.D., ET AL.: "Production of gamma-linolenic acid by transgenic plants expressing cyanobacterial or plant delta6-desaturase genes" PHYSIOL., BIOCHEM. MOL. BIOL. PLANT LIPIDS, [PROC. INT. SYMP. PLANT LIPIDS], 12TH (1997), 351-353. EDITOR(S): WILLIAMS, JOHN PETER; KHAN, MOBASHSHER UDDIN; LEM, NORA WAN. PUBLISHER: KLUWER, DORDRECHT, NETH. CODEN: 65BHAZ, XP002076486 see page 353	1-23
0,A	& BEREMAND, P.D. ET AL.: DISCLOSURE AT THE 12TH INTERNATIONAL SYMPOSIUM ON PLANT LIPIDS, HELD JULY 7-12, 1996, UNIVERSITY OF TORONTO,	1-23
A	TOEPFER R ET AL: "MODIFICATION OF PLANT LIPID SYNTHESIS" SCIENCE, vol. 268, 5 May 1995, pages 681-685, XP002014017 see the whole document	24,25
A	GIBSON, S., ET AL.: "Cloning of a temperature-regulated gene encoding a chloroplast omega-3 desaturase from Arabidopsis thaliana" PLANT PHYSIOLOGY, vol. 104, 1994, pages 1615-1621, XP002075189 see the whole document	25

.mation on patent family members

Interr nal Application No PC1/US 98/07178

	<del></del>		<del></del>	T	L	T
	ent document in search report		Publication date	Patent famil member(s)		Publication date
WO	9621022	A	11-07-1996	AU 4673 CA 2207 CN 1177 EP 0801	393 A 596 A 906 A 379 A 680 A 220 A	25-03-1997 24-07-1996 11-07-1996 25-03-1998 22-10-1997 04-08-1998
WO	9217580	A	15-10-1992	AU 1681 BR 9205 EP 0579 JP 6509	949 B 992 A 873 A 726 A 939 T 592 A	24-10-1996 02-11-1992 23-08-1994 26-01-1994 10-11-1994 01-10-1992
WO	9410189	A	11-05-1994	EP 0666	358 A 865 A 891 T	11-05-1994 16-08-1995 02-04-1996
WO	9113972	A	19-09-1991		722 A 595 A	04-03-1992 03-03-1998
WO	9211373	A	09 <b>-</b> 07-1992	AU 9116 DE 69113 DE 69113 DK 563 EP 0563		07-09-1995 22-07-1992 09-11-1995 13-06-1996 04-12-1995 06-10-1993 25-06-1996
WO	9311245	A	10-06-1993	AU 3228 CA 2124 EP 0616	923 B 893 A 673 A 644 A 701 T	27-02-1997 28-06-1993 10-06-1993 28-09-1994 23-02-1995
WO	9418337	A	18-08-1994		998 A 490 T	06-12-1995 16-07-1996
WO	.9411516	Α	26-05-1994		594 A 223 A	08-06-1994 26-05-1994

....ormation on patent family members

Interr I I I Application No PCT/US 98/07178

Patent document cited in search report		Publication date		atent family member(s)	Publication date		
WO 9411516	Α		EP JP	0668919 A 8503364 T	30-08-1995 16-04-1996		
W0 9606936	A	07-03-1996	AU BR CA EP HU JP PL	3410295 A 9509502 A 2198222 A 0778896 A 76842 A 10505237 T 319103 A	22-03-1996 30-09-1997 07-03-1996 18-06-1997 28-11-1997 26-05-1998 21-07-1997		